

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2001, 10:52:56 ; Search time 30.06 Seconds

(without alignments)  
33.883 Million cell updates/sec

Title: US-09-288-719-3

Sequence: 1 GGGSGGGRASGGGS 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	233	2 JC5322	p53 specific singl
2	68	84.0	268	2 A56446	Ig heavy chain V r
3	67	82.7	586	2 T26667	hypothetical prote
4	67	82.7	877	2 T43449	hypothetical prote
5	67	82.7	1433	2 A46053	bulbous pemphigoid
6	66	81.5	592	2 E82759	endc-1,4-delta-gluc
7	66	81.5	779	2 A35006	cell surface glyco
8	65	81.5	1585	2 T31611	hypothetical prote
9	65	80.2	104	2 JC4190	holotricin 3 precu
10	65	80.2	136	2 T29282	hypothetical prote
11	65	80.2	221	2 T04592	glycine-rich cell
12	65	80.2	316	2 T20497	hypothetical prote
13	65	80.2	643	1 KHRH2	keratin 1, type II
14	64	79.0	249	2 S41374	single chain Fv an
15	63	77.8	206	2 T53066	gene M-twist prote
16	63	77.8	214	1 KNNP25	glycine-rich prote
17	63	77.8	239	2 S49193	GCR 101 protein -
18	63	77.8	265	2 A40649	hypothetical prote
19	63	77.8	271	2 S34666	glycine-rich prote
20	63	77.8	312	2 T25048	hypothetical prote
21	63	77.8	431	1 WIRH2G	homeotic protein H
22	63	77.8	433	2 S20963	homeotic protein H
23	63	77.8	1454	2 T13709	diacylglycerol kin
24	62	76.5	80	2 T10550	hypothetical prote
25	62	76.5	157	1 S14857	glycine-rich prote
26	62	76.5	157	1 S04536	embryonic abundant
27	62	76.5	165	1 KNR2G1	glycine-rich cell
28	62	76.5	168	1 S12312	glycine-rich RNA-b
29	62	76.5	207	2 T07381	glycine-rich prote

30	62	76.5	280	2 A42424	chitinase (EC 3.2.
31	62	76.5	291	1 S31415	glycine-rich prote
32	62	76.5	322	2 A45036	single-stranded-DN
33	62	76.5	333	2 S16318	homeotic protein H
34	62	76.5	378	2 S04336	U1 snRNP 70K prote
35	62	76.5	440	2 S71795	transcription fact
36	62	76.5	495	1 S31223	transcription fact
37	62	76.5	528	2 G02127	transcription fact
38	62	76.5	569	1 KRMS1	fus-like protein -
39	62	76.5	570	2 S07330	keratin, 59K type
40	62	76.5	979	2 A35913	regulatory factor
41	62	76.5	1969	2 T08875	histidine kinase h
42	62	76.5	2783	1 A41948	alpha-fetoprotein
43	61	75.3	102	2 D71415	hypothetical prote
44	61	75.3	165	2 T03583	glycine-rich RNA-b
45	61	75.3	316	1 A38743	loricrin - human

#### ALIGNMENTS

RESULT 1  
JC5322  
p53 specific single-chain antibody Pab421 - human  
C:Species: Homo sapiens (man)  
C:Date: 15-May-1997 #sequence\_revision 15-May-1997 #text\_change 18-Jul-1997  
C:Accession: JC5322  
R:Janot, C.B.; Hynes, N.E.  
Biochem. Biophys. Res. Commun. 230, 242-246, 1997  
A>Title: Characterization of scFv-421, a single-chain antibody targeted to p53.  
A:Reference number: JC5322; MIMD:97168950  
A:Accession: JC5322  
A:Molecule type: mRNA  
A:Residues: 1-233 <JN>  
A:Experimental source: hybridoma cell  
C:Comment: This protein specifically binds the tumor suppressor protein p53. It resto

Query Match 100.0%; Score 81; DB 2; Length 233;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGSGGGRASGGGS 15  
DB 110 GGGSGGGRASGGGS 124

RESULT 2  
A56446  
Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)  
C:Species: Mus musculus (house mouse)  
C:Date: 19-Jan-1996 #sequence\_revision 19-Jan-1996 #text\_change 16-Aug-1996  
C:Accession: A56446  
R:Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.  
J. Biol. Chem. 270, 7829-7835, 1995  
A>Title: A high affinity digoxin-binding protein displayed on M13 is functionally ide  
A:Reference number: A56446; MIMD:95229583  
A:Accession: A56446  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-268 <TN>  
A:Cross-references: GB:U20617  
C:Keywords: heterotrimer; immunoglobulin

Query Match 84.0%; Score 68; DB 2; Length 268;  
Best Local Similarity 86.7%; Pred. No. 0.53;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGSGGGRASGGGS 15  
DB 121 GGGSGGGRASGGGS 135

RESULT 3  
726667  
hypothetical protein Y38E10A.g - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T26667  
R:Wallis, J.

Submitted to the EMBL Data Library, September 1999

A:Reference number: Z20252

A:Accession: T26667

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-586 <MIT>

A:Cross-references: EMBL:AL110484; NID:e1542205; PIDN:CAB54408.1; CESP:Y38E10A.g

A:Experimental source: clone Y38E10A

C:Genetics:

A:Gene: CESP:Y38E10A.g

A:Introns: 105/2; 174/1; 248/1; 372/3; 436/3

Query Match 82.7%; Score 67; DB 2; Length 586;  
Best Local Similarity 80.0%; Pred. No. 1.3;

Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGGSGGRASGGGGS 15  
||||| :|||||  
Db 461 GGGGSGGYCAGGCGS 475

RESULT 4

T43449  
hypothetical protein DKFZP434P0917.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000

C:Accession: T43449

R:Boecker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: Z22513

A:Accession: T43449

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-877 <AA>

A:Cross-references: EMBL:AL133612

A:Experimental source: adult testis; clone DKFZP434P0917

C:Genetics:

A:Note: DKFZP434P0917.1

Query Match 82.7%; Score 67; DB 2; Length 877;  
Best Local Similarity 80.0%; Pred. No. 1.8;

Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGGSGGRASGGGGS 15  
||||| :|||||  
Db 143 GGGGSGGSSGGGGS 157

RESULT 5

A46053  
bullous pemphigoid antigen, BPAG2, type XVII collagen alpha 1-chain - mouse

C:Species: Mus musculus (house mouse)

C:Date: 21-Sep-1993 #sequence\_revision 21-Sep-1993 #text\_change 05-Nov-1999

C:Accession: A46053

R:Li, K.; Tamai, K.; Tan, E.M.L.; Uitto, J.

J. Biol. Chem. 268, 8825-8834, 1993

A:Title: Cloning of type XVII collagen. Complementary and genomic DNA sequences of mouse

egment, and unusual features in the 5'-end of the gene and the 3'-untranslated region of

A:Reference number: A46053; MUID:93232041

A:Accession: A46053

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1433 <L11>

A:Cross-references: GB:L08407; NID:g309182; PIDN:AAA37443.1; PID:g309183  
A:Note: sequence extracted from NCBI backbone (NCBIN:129627, NCBI:P:129628)

Query Match 82.7%; Score 67; DB 2; Length 1433;  
Best Local Similarity 80.0%; Pred. No. 2.8;

Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGGSGGRASGGGGS 15  
||||| :|||||  
Db 443 GGGGSGGASGGGGA 457

RESULT 6

E82759  
endo-1,4-beta-glucanase XF0818 (imported) - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: E82759

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: E82759

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-592 <SIM>

A:Cross-references: GB:AE003921; GB:AE003849; NID:g9105710; PIDN:AAF83628.1; GSPDB:GN

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alverenga, R.

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carriro, D.M.; Carrer

as-Neto, E.; Docena, C.; El-Porty, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr

J.D.; Jungueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurmae, E.E.; La

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeti, C.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa

M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF0818

Query Match 81.5%; Score 66; DB 2; Length 592;  
Best Local Similarity 85.7%; Pred. No. 1.7;

Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGGSGGRASGGGGS 14  
||||| :|||||  
Db 467 GGGGSGGSGGGGG 480

RESULT 7

A35006  
cell surface glycoprotein msp130 precursor - sea urchin (Strongylocentrotus purpuratu

C:Species: Strongylocentrotus purpuratus (purple urchin)

C:Date: 20-Jul-1990 #sequence\_revision 20-Jul-1990 #text\_change 09-Sep-1997

C:Accession: A35006; A45939

R:Part, B.A.; Parks, A.L.; Raff, R.A.

J. Biol. Chem. 265, 1408-1413, 1990

A:Title: Promoter structure and protein sequence of msp130, a lipid-anchored sea urch

A:Reference number: A35006; MUID:90110195

A:Accession: A35006

A:Status: preliminary

A:Molecule type: DNA; mRNA

A:Residues: 1-779 <PAR>

A:Cross-references: GB:M31751; NID:g161543; PID:g161544; GB:J05204

R:Leat, D.S.; Anstrom, J.A.; Chin, J.E.; Harkey, M.A.; Showman, R.M.; Raff, R.A.

Dev. Biol. 121, 29-40, 1987  
A:Title: Antibodies to a fusion protein identify a cDNA clone encoding msp130, a primary

A:Reference number: A45939; MID:87191419

A:Accession: A45939

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 542-779 <LEA>

A:Cross-references: GB:M16457; NID:g161539; PID:g161540

C:Keywords: glycoprotein

Query Match 81.5%; Score 66; DB 2; Length 779;  
Best Local Similarity 78.6%; Pred. No. 2.1;

Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGSGGRASGGG 14  
||||| ||| ||| |||  
Db 57 GGGGAGGRGGGGG 70

RESULT 8

TJ1611

hypothetical protein Y508A.g - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: TJ1611

R:Steward, C.  
submitted to the EMBL Data Library, September 1999

A:Reference number: Z21047

A:Accession: TJ1611

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1585 <MTL>

A:Cross-references: EMBL:AL117200; NID:e1549770; PID:CAE55050.1; CESP:Y508A.g

A:Experimental source: clone Y508A

C:Genetics:

A:Gene: CESP:Y508A.g

A:Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match 81.5%; Score 66; DB 2; Length 1585;  
Best Local Similarity 85.7%; Pred. No. 3.9;

Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGSGGRASGGG 14  
||||| ||| ||| |||  
Db 452 GGGGSGGVSAGGG 465

RESULT 9

JC4190

holotrichin 3 precursor - Holotrichia diomphalia

N:Alternate names: antifungal protein

C:Species: Holotrichia diomphalia

C>Date: 04-Oct-1995 #sequence\_revision 10-Nov-1995 #text\_change 05-Nov-1999

C:Accession: JC4190

R:Lee, S.Y.; Moon, H.J.; Kurata, S.; Natori, S.; Lee, B.L.

Biol. Pharm. Bull. 18, 1049-1052, 1995

A:Title: Purification and cDNA cloning of an antifungal protein from the hemolymph of Hc

A:Reference number: JC4190; MID:96073722

A:Accession: JC4190

A:Molecule type: mRNA

A:Residues: 1-104 <LEE>

A:Cross-references: DBJ:DJ3744; NID:g1088433; PID:BA02889.1; PID:di003394; PID:g17861

C:Comment: This protein is a Gly- and His-rich protein and a constitutive protein of lat

C:Keywords: hemolymph

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-104/Product: holotrichin 3 #status predicted <MAT>

Query Match 80.2%; Score 65; DB 2; Length 104;

Best Local Similarity 80.0%; Pred. No. 0.51;  
Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGSGGRASGGG 15  
||||| ||| ||| ||| |||  
Db 64 GGGGRGGGGSGGG 78

RESULT 10

T29282

hypothetical protein C34D4.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000

C:Accession: T29282

R:Du, Z.; Le, T.T.  
submitted to the EMBL Data Library, May 1996

A:Description: The sequence of C. elegans cosmid C34D4.

A:Reference number: Z20500

A:Accession: T29282

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-136 <DUZ>

A:Cross-references: EMBL:U58755; PID:AB00696.1; GSPDB:GN00022; CESP:C34D4.11

A:Experimental source: strain Bristol N2; clone C34D4

C:Genetics:

A:Gene: CESP:C34D4.11

A:Map position: 4

A:Introns: 20/1; 66/1; 98/1; 116/1

C:Superfamily: Arabidopsis glycine-rich protein 3

Query Match 80.2%; Score 65; DB 2; Length 136;  
Best Local Similarity 78.6%; Pred. No. 0.64;

Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGSGGRASGGG 14  
||||| ||| ||| |||  
Db 97 GGGGSGGRGGGGG 110

RESULT 11

T04592

glycine-rich cell wall structural protein homolog F23E13.120 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 14-May-1999

C:Accession: T04592

R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Jesse,

submitted to the Protein Sequence Database, March 1998

A:Reference number: Z15378

A:Accession: T04592

A:Molecule type: DNA

A:Residues: 1-221 <BEV>

A:Cross-references: EMBL:AL022141

C:Experimental source: cultivar Columbia; BAC clone F23E13

C:Genetics:

A:Map position: 4

A:Note: F23E13.120

Query Match 80.2%; Score 65; DB 2; Length 221;  
Best Local Similarity 78.6%; Pred. No. 0.96;

Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGSGGRASGGG 14  
||||| ||| ||| |||  
Db 98 GGGGSGGSGGGG 111

RESULT 12

T20497

hypothetical protein F02D10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T20497

R:Swingburne, J.



Query Match 77.8%; Score 63; DB 2; Length 206;  
Best Local Similarity 73.3%; Pred. No 1.5;  
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGSGGRASGGGS 15  
|||||  
Db 83 GGGGAGGGGGGGGS 97

Search completed: March 15, 2001, 10:52:57  
job time: 1006 sec